

SEQ ID NO:1

Tc	MRKSVC PKQKFFFSAFPFFFFFCVFPLISRTGQEKLLFDQYKIIKGEKKKXKNQRANRREHQKREIMRFKNS	75
Tc	FTCIDMHTEGEAARIVTSGLPHIPGSNMAEKKAYLQENMDYLRRGINLEPRGHDDMEGAFLEFDPIIEGADLGXVF	150
Tc	MDTGGYLNMCQHNSIAAVTAAVETGIVSVPAKATNVPVVLDTAGLVRGTAHLQSGTESEVSNASIIINVPSFLYQ	225
Tc	QDVVVVLPKPYGEVRVDIAFGGNFFAIVPAEQLGIDISVQNLSRLQEAGELLRTEINRSVKVQHPQLPHINTVDC	300
Tc	VEIYGPPTNPEANYKNVVIFGNRQADR SPCGT GTSAKMATLYAKGQLRIGETFVYESILGSLFQGRV--LGEE	371
Tc	RIPGVKVPVTKDAEEGMLVVTAETGKAFINGFNTMLFDPTDPFKNGFTLKQ*	423

SEQ ID NO:2

Tc	RTGQEKLLFDQKYKIIKGEKKEKKKNQRANRREHQQKREIMRFKKS	75
Tc	FTCIDMHTEGEAARI <u>VTSGLP</u> HIPGSNMAEKKAYLQENMDYLRRGIMLEPRGHDDMFGAFLFDPISEGADLGMVF	150
Tc	MDTGGYLNMCGHNSIAAVTAAVETGIVSVPAKATNVPVVLDT <u>PAGLV</u> RGT <u>TAHLQ</u> SGTESEVSNASIINVPSFLYQ	225
Tc	QDVVVVLPKPYGEVR <u>VDIAFGGN</u> FFAIVPAEQLGIDISVQNLRLQEAGELLRTEINRSVKVQHPQLPHINTVDC	300
Tc	VEIYGPPTNPEANYKNVVIFGNROADR SPCGT GTSAKMATLYAKGQLRIGETFVYESILGSLFQGRV--LGEE	371
Tc	RIPGVKVPVTKDAEEGMLVVTAETGKAFIMGFNTMLFDPTDPFKNGFTLKQ*	423

SEQ ID NO:3

Tc MRKSVCPKQKFFFSAFPFFFFCVFPLIS

SEQUENCE ID NO.4

73
150
225
300
371
423

SEQUENCE ID NO:5

Cs	-----MKFSKG	6
Cs	IHAIDSHTMGEPTRIVVGGIPQINGETMADKKKYLEDNLDYVRTALMHEPRGHNDMFGSIITSSNNKEADFGIIF	31
Cs	MDGGGYLNMCGHGSIGAAATVAVETGMVEMVEPVTNIN--MEAPAGLIKAKVMVEN---EKVKEVSITNVPSFLYM	151
Cs	EDAKLEVPSLNKTITFDISFGGSFFAIHAKELGVKVETSQVDVLKKGIEIRDLINERIKVQHPELEHKTVDL	226
Cs	VEIYDEPSNPEATYKNVVIFGQGQVDR SPCGT GTSAKLATLYKKGHLKIDEKEVYESITGTMEKGRV--LEET	297
Cs	KVGEFD-----AIIPEITGGAYITGENHEVIDPEDPLKYGFTV*--	335

SEQ ID NO:6

Pa	-----MQR	3
Pa	IRIIDSHTGGEPTRLVIGGFDPDLGQGDMABRRRLGERHDAWRAACILEPRGSDVLVGALLCAPVDPEACAGVIF	78
Pa	FNNSGYLGMCGHGTIGLVASLAHLGRIGPGV-----HRIETPVGEVEATLH-----EDGSVSVRNVPAYRYR	140
Pa	RQVSVEVPGI-GRVSGDIAGGNWFFLVAGH--GQRLAGDNLDALTAYTVAVQQALDD---QDIRGEDGGAIDH	208
Pa	IELFAD--DPHADSRNFVLCPGKAYDR SPCGT GTSAKLACLAADGKLLPGQPWRQASVIGSQFEGRYEWLOGQ	279
Pa	PGGPVPTIRGRAHVSAEATLLLADDDPFAWGIRR* -----	314

SEQ ID NO: 7

Polypyrimidine rich region



Splice leader
acceptor sites

Signal Peptide

CCTTTTCTTTTTAAAAACAAAAAAATTCCGGGGGAATATGGAACAGGGTATATGCGTAAAAGTGTCTGTCCCAACAAAAATTTTTT 90
TTTTCCGCCTTCCCATTTTTTTTTTTTTTTTGTGTGTTTCCCTTGATCTCTCGAACAGGGCAGGAAAAGCTTCTGTTTGACCAAAAATAT 120
F S A F P F F F F F C V F P L I S R T G Q E K L L F D Q K Y 42
AAAAATTATTAAGGGCGAGAAAAAGAAAAAGAAAAAATCAACGAGCAAAACAGGAGAGAACACCAACAAAAAGGGAAATTATGCGATT 270
K I I K G E K K E K K K N Q R A N R R E H Q Q K R E I M R F 72
AAGAAATCATTACATGCATCGACATGCATACGGAAGGTGAAGCAGCACGGATTGTGACGAGTGGTTTGCCACACATTCCAGGTTGCAAT 360
K K S F T C I D M H T E G E A A R I V T S G L P H I P G S N 102
ATGGCGGAGAAGAAAGCATACCTGCAGGAAAACATGGATTATTTGAGCGGTGGCATAATGCTGGAACACGTTGGTTCATGATGATATGTTT 430
M A E K K A Y L Q E N M D Y L R R G I M L E P R G H D D M F 132
GGAGCCTTTTTATTTGACCCTATTGAAGAAGGCGCTGACTTGGGCATGGTATTCATGGATACCGGTGGCTATTTAAATATGTGTGGACAT 520
G A F L F D P I E E G A D L G M V F M D T G G Y L N M C G H 162
AACTCAATTGCAAGCGTTACGGCGGCAGTTGAAACGGGAATTGTGAGCGTGCCGCGCAAGGCAACAAATGTTCCGGTTGTCTGGACACA 610
N S I A A V T A A V E T G I V S V P A K A T N V P V V L D T 192
CCTGCGGGGTTGGTGCAGCGGTACGGCACACCTTCAGAGTGGTACTGAGAGTGAAGGTGTCAAATGCGAGTATTATCAATGTACCTTCATT 700
P A G L V R G T A H L Q S G T E S E V S N A S I I N V P S F 222
TTGTATCAGCAGGATGTGGTGGTTGTGTGGCCAAAGCCCTATGGTGAAGTACGSGTTGATATTGCATTGGAGGCAATTTTTTCCGCATT 790
L Y Q Q D V V V V L P K P Y G E V R V D I A F G G N F F A I 252
GTTCCCGCGGAGCAGTTGGGAATTGATATCTCCGTTCAAACCTCTCCAGGCTGCAGGAGGCAGGAGAACTTCTGCGTACTGAAATCAAT 330
V P A E Q L G I D I S V Q N L S R L Q E A G E L L R T E I N 282
CSCAGTGTGAAGGTTCAGCACCTCAGCTGCCCCATATTAACTGTGGACTGTGTGAGATATACGGTCCGCCAACGAACCCGGAGGCA 970
R S V K V Q H P Q L P H I N T V D C V E I Y G P P T N P E A 312
AACTACAAGAACGTTGTGATATTGGCAATCGCCAGGCGGATCGCTCTCCATGTGGGACAGGCACCAGCGCCCAAGATGGCAACACTTTAT 1060
N Y K N V V I F G N R Q A D R S P C G T G T S A K M A T L Y 342
GCCAAAGGCCAGCTTCGCATCGGAGAGACTTTTGTGTACGAGACATACTCGGCTCACTCTTCCAGGSCAGGGTACTTGGGAGGAGCGA 1150
A K G Q L R I G E T F V Y E S I L G S L F Q G R V L G E E R 372
ATACCGGGGTTGAAGGTGCCGTTGACCAAGATGCCGAGGAAGGGATGCTCGTTGTAAACGGCAGAAATTACTGGAAGGCTTTTATCATG 1240
I P G V K V P V T K D A E E G M L V V T A E I T G K A F I M 402
GGTTTCAACACCATGCTGTTTGACCAACGGATCGGTTTAAGAACGGATTCAATTAAAGCAGTAGATCTGGTAGAGCAGACAACTATT 1330
G F N T M L F D P T D P F K N G F T L K Q 423
GGGGAACAGTGCAGAACAGGTGCTGCTACGTGAAGGATATTGAATGAATCGTTTTTTTTTATTTTTTATTTTTTATTAGTGCATT 1420
ATTATTAATTTTTTTTTTTTGTGTTTGGGGTTTCAACGGTACCGCGTGGGAGCAGGSAAGCGATAGCGGCCGACAAATTTTTTGTCTTTAT 1510
TTTCATTTTCATCTTCCTACCCAACCCCTTGGTTCCACCGGTGCGGGCGGGGCTTGTGTGGTGGAGGAGTCTAAATCCCGCACCTCGG 1600
AGGAATAAACATATTTCAATTTTCATATCTTGAATCAAAAGGCAT 1651

Polyadenylation site

Obs : Underlined the sequenced peptides used to deduce degenerated primers for cloning

(d) Nucleotide sequence and peptide sequence TcPA45

SEQ ID NO.8

TTTTCCGCTTCCCATTTTTTTTTTTTTTTTGTGTGTTCCCTTCTCTCTG M R K S V C P K Q K F F 130
F S A F P F F F F F F C V F P L I S R T G Q E K L L F D Q K Y 42
AAAATTATTAAGGGCGAGAAAAAAGAAAAAATCAACGAGCAAAACAGGAGAGAACACCAACAAAAAGGGAAATTATGCGATTT 270
K I I K G E K K E K K K N Q R A N R R E H Q Q K R E I M R F 72
AAGAAATCATTACATGCATCGACATGCATACGGAAGGTGAAGCAGCAGGATTGTGACGAGTGGTTTGGCACACATTCCAGGTTCSAAT 350
K K S F T C I D M H T E G E A A R I V T S G L P H I P G S N 102
ATGGCGGAGAAGAAAGCATACCTGCAGGAAAAACATGGATTATTTGAGGCGTGGCATAATGCTGGAACACAGTGGTCATGATGATATGTTT 430
M A E K K A Y L Q E N M D Y L R R G I M L E P R G H D D M F 132
GGAGCCTTTTTATTTGACCCTATTGAAGAAGGCGCTGACTTGGGCATGGTATTCATGGATACCGGTGGCTATTTAAATATGTGTGGACAT 520
G A F L F D P I E E G A D L G M V F M D T G G Y L N M C G H 162
AACTCAATTGCAGCGGTTACGGCGGCGATTGAAACGGGAATTGTGAGCGTGCCGGCGAAGGCAACAAATGTTCCGGTTGTCTGGACACA 610
N S I A A V T A A V E T G I V S V P A K A T N V P V L D T 192
CCTGCGGGGTTGGTGC GCGGTACGGCACACCTTCAGAGTGGTACTGAGAGTGAGGTGTCAATGCGAGTATTATCAATGTACCCCTCATTT 700
P A G L V R G T A H L Q S G T E S E V S N A S I I N V P S F 222
TTGTATCAGCAGGATGTGGTGGTTGTGTGCCAAAGCCCTATGGTGAAGTACGGGTTGATATTGCATTTGGAGGCAATTTTTTCGCCATT 750
L Y Q Q D V V V V L P K P Y G E V R V D I A F G G N E F A I 252
GTTCCCGCGGAGCAGTTGGGAATTGATATCTCCGTTCAAAAACCTCTCCAGGCTGCAGGAGGCAGGAGAACTTCTGCSTACTGAAATCAAT 550
V P A E Q L G I D I S V Q N L S R L Q E A G E L L R T E I N 292
CGCAGTGTGAAGGTTGAGCACCCTCAGCTGCCCATATTAACACTGTGGACTGTGTTGAGATATACGCTCCGCCAACGAACCCGGAGGCA 970
R S V K V Q H P Q L P H I N T V D C V E I Y G P P T N P E A 312
AACTACAAGAACGTTGTGATATTTGGCAATCGCCAGGCGGATCGCTCTCCATGTGGGACAGGCACACAGCGCCAGATGGCAACACTTTAT 1060
N Y K N V V I E G N R Q A D R S P C G T G T S A K M A T L Y 342
GCCAAAGGCCAGCTTCGCATCGGACAGACTTTTGTGTACGAGAGCATACTCGGCTCACTCTTCCAGGGCAGGGTACTTGGGGAGGAGCGA 1150
A K G Q L R I G E T F V Y E S I L G S L F Q G R V L G E E F 372
ATACCGGGGTGAAGGTGCCGGTGACCAAAGATGCCGAGGAAGGGATGCTCGTTGTAACGGCAGAAATTACTGGAAAGGCTTTTATCATG 1240
I P G V K V P V T K D A E E G M L V V T A E I T G K A F I M 402
GTTTTCAACACCATGCTGTTTGACCAACGGATCCGTTTAAAGAACGGATTACATTAAAGCAGTAGATCTGGTAGAGCACAGAACTATT 1330
G F N T M L F D P T D P F K N G F T L K Q * 423
GGGGAACACGTCGGAACAGGTGCTGCTACGTGAAGGTATTGAATGAATCGTTTTTTTTTTATTTTTTTATTTTTTTATTAGTGCATT 1420
ATTATTAATTTTTTTTTTTTGTGTTTGGGGTTTCAACGGTACCGCGTTTGGAGCAGGGAAGCGATAGCGGCCGGACAATTTTTTGTGTTTTAT 1510
TTTCATTTTCATCTTCTACCCAAACCCCTTGGTTCCACCGGTGCGCGCGGGTCTTGTGGGTGAGGAGTCCTAAATCCCGCACCTCGG 1600
AGGAATAAACATATTTCAATTTTCATATCTTGAATCAAAAGGCAT 1651

Polyadenilation site

Obs : Underlined the sequenced peptides used to deduce degenerated primers for cloning

(a) Nucleotide sequence and peptide sequence TcPA45

SEQ ID NO.9

F S A F P F F F F F C V F P L I S R T G Q E K L L F D Q K Y 270
 AAAATTATTAAGGGCGAGAAAAAGAAAAAATCAACGAGCAACAGGAGAGAACCACCAACAAAAAGGAAATTATGCGATTT 360
K I I K G E K K E K K K N Q R A N R R E H Q Q K R E I M R F 102
 AAGAAATCATTACATGCATCGACATGCATACGGAAGGTGAAGCAGCAGGATTGTGACGAGTGGTTTGCCACACATTCAGGTTGCAAT 430
K K S F T C I D M H T E G E A A R I V T S G L P H I P G S N 132
 ATGGCGGAGAAGAAAGCATACCTGCAGGAAAACATGGATTATTTGAGGCGTGGCATAATGCTGGAACCACGTTGGTCATGATGATATGTTT 520
 M A E K K A Y L Q E N M D Y L R R G I M L E P R G H D D M F 162
 GGAGCCTTTTATTTGACCCTATTGAAGAAGGCGCTGACTTGGGCATGGTATTTCATGGATAACGGTGGCTATTTAAATATGTGTGGACAT 610
 G A F L F D P I E E G A D L G M V F M D T G G Y L N M C G H 192
 AACTCAATTGCAGCGGTTACGGCGGCAGTTGAAACGGGAATTGTGAGCGTGCCGGCGAAGGCAACAAATGTTCCGGTTGTCTCGGACACA 700
 N S I A A V T A A V E T G I V S V P A K A T N V P V V L D T 222
 CCTGCGGGGTTGGTGCGCGGTACGGCACACCTTCAGAGTGGTACTGAGAGTGAGGTGTCAAATGCGAGTATTATCAATGTACCTCATTTT 790
P A G L V R G T A H L Q S G T E S E V S N A S I I N V P S F 252
 TTGTATCAGCAGGATGTGGTGTGTGTTGCCAAAGCCCTATGTTGAAGTACGGGTTGATATTGCATTTGGAGGCAATTTTTTCGCCATT 380
L Y Q Q D V V V V L P K P Y G E V R V D I A F G G N F F A I 282
 GTTCCCGCGGAGCAGTTGGGAATTGATATCTCCGTTCAAACCTCTCCAGGCTGCAGGAGGCAGGAGAACTTCTGCGTACTGAAATCAAT 970
 V P A E Q L G I D I S V Q N L S R L Q E A G E L L R T E I N 312
 CGCAGTGTGAAGGTTACGACCCCTCAGCTGCCCATATTAACACTGTGGACTGTGTGAGATATACGGTCCGCCAACGAACCCGGAGGCA 1060
 R S V K V Q H P Q L P H I N T V D C V E I Y G P P T N P E A 342
 AACTACAAGAAGCTTGTGATATTTGGCAATCGCCAGGCGGATCGCTCTCCATGTGGGACAGGCACCGCCCAAGATGGCAACACTTTAT 1150
 N Y K N V V I F G N R Q A D R S P C G T G T S A K M A T L Y 372
 GCCAAAGGCCAGCTTCGCATCGGAGAGACTTTTGTGTACGAGAGCATACTCGGCTCACTCTTCCAGGGCAGGGTACTTGGGGAGGAGCGA 1240
A K G Q L R I G E T F V Y E S I L G S L F Q G R V L G E E R 402
 ATACCGGGGGTGAAGGTGCCGGTGACCAAAGATGCCGAGGAAGGGATGCTCGTTGTAACGGCAGAAATTACTGGAAGGCTTTTATCATG 1330
 I P G V K V P V T K D A E E G M L V V T A E I T G K A F I M 423
 GGTTCACACCATGCTGTTTGACCCAACGGATCCGTTTAAGAACGGATTACATTAAAGCAGTAGATCTGGTAGAGCACAGAACTATT 1420
 G F N T M L F D P T D P F K N G F T L K Q *
 GGGGAACACGTSCGAACAGGTGCTGCTACGTGAAGGGTATTGAATGAATCGTTTTTTTTTATTTTTTATTTTTTATTAGTGCATT 1510
 ATTATTAATTTTTTTTTTTGTTTTGGGGTTTCAACGGTACCGCGTTGGGAGCAGGGAACGATAGCGGCCGACAAATTTTTTGCTTTTTAT 1600
 TTTTATTTTATCTTCCCTACCCCAACCCCTTGGTTCCACCGGTGCGGGCGGGGTCCTTGTGGGTGGAGGAGTCTTAAATCCCGCACCTCGG 1651
 AGGAATAAACATATTTCAATTTTCATATCTTGAATCAAAAGGCAT

Polyadenilation site

Obs : Underlined the sequenced peptides used to deduce degenerated primers for cloning

(b) Nucleotide sequence and peptide sequence TcPA45

SEQUENCE ID NO:10

Signal peptide

1
M R K S V C P K Q K F F

Nucleotide sequence of signal sequence TcPA45

SEQUENCE ID NO:11

ATGCGATT<

	K R F	340
.....		333
X K S I T C T O M R T X C I A A X V T S Q L P X I P O S X		433
.....		337
M A Z X K A Y L Q Z X K O T L A I C T M L E P L O X O O M P		320
.....		342
O A T L P O P I Z Z C A O L C M Y P M O T C C T L M K C C X		610
.....		332
M S C A A Y T A A Y S T C T Y S P A X A T V V P V V L O T		703
.....		333
P L C L V I C T A Y L Q S C T C C S V S W A S T T R V P S T		730
.....		332
L T O Q O Y V V Y L P X P T C S T A V O T J P C Q M P T A S		640
.....		342
V P A Z Q L C C C S Y Q V L S L L Q T A C X L L R T S C K		970
.....		332
R S Y Z V Q K P Q L P X I K T V O C Y C T C P P T K P S A		1040
.....		342
M T X M V V T P C M R Q A O R S P C C T C T S A X M A T L Y		1150
.....		372
L X C O L I C C T T Y T S S L L O S L T Q C R V L C E S R		1240
.....		402
I P C Y X Y P Y T X O A S C M L V Y T A S I T C X A P S M		1330
.....		430
C T X T K L T O P T O P P Y C P T T L K O		1420
.....		3520
.....		1400
.....		1452

SEQ ID NO : 12

5'TTICCRAADATIACIACGTT 3'

SEQ ID NO : 13

5' ATHGCITTYGGIGGIAAYTTT 3'

SEQ ID NO : 14

5' TTICCRAADATIACIACGTT 3'

SEQ ID NO : 15

5' CTCTCCCATGGGGCAGGAAAAGCTTCTG 3'

SEQ ID NO : 16

5' CTGAGCTCGACCAGATCTATCTGC 3'

SEQ ID NO : 17

```

1 ccttttttctt tttaaaaaca aaaaaaatcc cgggggggaat atggaacagg gtatatgogt
61 aaaagtgtct gtcccaaaca aaaatttttt tttccgoot tcccattttt tttttttttt
121 tgttgtgttc ccttgatctc tcgaacaggg caggaaaagc ttctgtttga ccasaaatat
131 aaaattatta agggcgagaa aaaaagaaaag aaaaaaaatc aacgagcaaa caggagagaa
241 caccaacaaa aaagggaat tatgcgattt aagaaatcat tcacatgcat cgacatgcat
301 acggaaggtg aagcagcacg gattgtgacg agtggtttgc cacacattcc aggttcgaat
361 atggcgagaa agaaagcata cctgcaggaa aacatggatt atttgaggog tggcataatg
421 ctggaaccac gtggtcatga tgatatgttt ggagcctttt tatttgacct tattgaagaa
431 ggcgctgact tgggcatggt attcatggat accggtggct atttaaatat gtgtggacat
541 aactcaattg cagcggttac ggcggcagtt gaaacgggaa ttgtgagcgt gccggcgag
601 gcaacaaatg ttccggttgt cctggacaca cctgcggggg tgggtgcgcg tacggcacac
661 cttcagagtg gtactgagag tgagggtgtc aatgcgagta ttatcaatgt accctcatct
721 ttgtatcagc aggatgtggt ggttctgttg ccaaagccct atggtgaagt accggttgat
731 attgcatttg gaggcaattt tttcgccatt gtcccgcggt agcagttggg aattgatato
841 tccgttcaaa acctctccag gctgcaggag gcaggagaac ttctgogtac tgaaatcaat
901 cgcagtgtga aggttcagca ccttcagctg ccccatatta acaactgtga ctgtgttgag
961 atatacggtc cgccaacgaa ccgggaggca aactacaaga acgttgtgat atttggcaat
1021 cgccaggcgg atcgctctcc atgtgggaca ggcaccagcg ccaagatggc aacactttat
1081 gccaaaggcc agcttcgcat cggagagact tttgtgtacg agagcatact cggctcaactc
1141 ttccagggca gggtaacttg ggaggagcga ataccggggg tgaagggtgc ggtgaocaaa
1201 gatgocgagg aagggatgct cgttctaacg gcagaaatta ctggaaaggc ttttatcatg
1261 ggtttcaaca ccattgctgt tgacccaacg gatccgttta agaacggatt cacattaaag
1321 cagtagatct ggtagagcac agaaactatt ggggaacacg tgcgaacagg tgcgtgtaag
1381 tgaagggtat tgaatgaatc gttttttttt attttttttt tttattttta ttagtgcatt
1441 attattaaat tttttttttg ttttgggggt tcaacgggtc cgcgttggga gcagggaagc
1501 gatagcggcc ggacaatttt ttgcttttat tttcattttt atcttcttac ccaacccctc
1561 tggttccacc ggtcgcggcg gggctctgtg ggtggaggag tctaaaatcc cgcacctcgg
1621 aggaataaac atattttcaat ttcatatctt ggaatcaaaa ggcat

```

SEQ ID NO : 18

WIIK

SEQ ID NO : 19

IVTGSLPDISG

SEQ ID NO : 20

ATNVPVVLDTAGLVR

SEQ ID NO : 21

VDIAFGGNF

SEQ ID NO : 22

NVVIFGNR

SEQ ID NO : 23

MATLYAK

SEQ ID NO : 24

5' TCCGTATCCATGTCGATGC 3'

SEQ ID NO : 25

5' TATTATTGATACAGTTTCTG 3'

SEQ ID NO : 26

5' CTCTCCCATGGGGCAGGAAAAGCTTCTG 3'